Predicting What We Breathe

Jeanne Holm, Mohammad Pourhomayoun, Dawn Comer
Kabir Nagrecha, Pratyush Muthukumar

City of Los Angeles
California State University Los Angeles
AI Agora
Open AQ

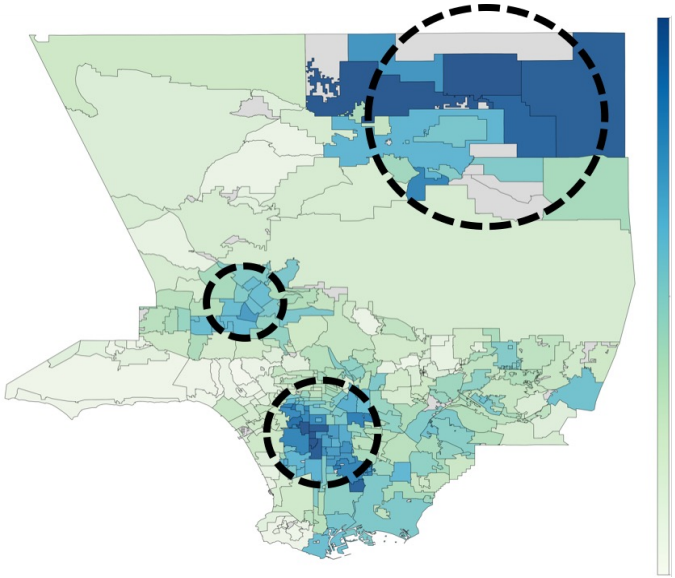
Air Pollution

- Urban Air pollution is mostly a man-maid problem. It is responsible for the early deaths of 7 million people every year. It means that every 5 seconds, somebody around the world dies prematurely from the effects of air pollution.
- Unfortunately, the minority and low-income communities face higher exposure to air pollutants and experience greater health impacts.

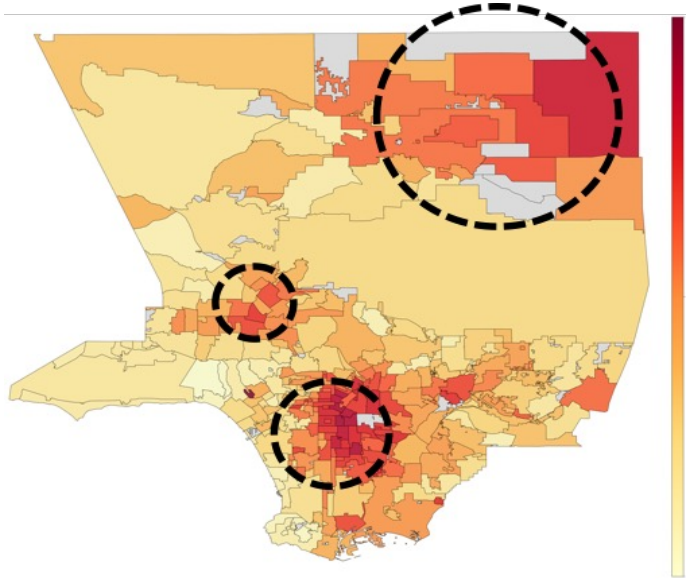
[1] UN Report 2019, UNICEF, https://news.un.org/en/story/2019/06/1039661
[2] The American Lung Association, “Disparities in the Impact of Air Pollution”

Air Pollution

Our study on the people of California demonstrates that minority and low-income communities tend to be exposed to higher levels of air pollution and hit hardest by the adverse health consequences of air pollution



Left: Asthma Rate By Neighborhood



Right: Poverty Percentage By Neighborhood

K. Marlis, J. West, D. Comer, I. Burga, J. Taub, C. F. Calvert, J. Holm, and M. Pourhomayoun, “A Comprehensive Analysis of Air Pollution and Equity During COVID-19 in Los Angeles County,” The 17th International Conference on Data Science, ICDATA'21: July 26-29, 2021

Air Pollution

- The first and the most important step in mitigating the air pollution risks is to understand the sources of it, discover the patterns, and predict it in advance.
- By enhancing human understanding and prediction of air quality, local governments, health providers, and others can help mitigate the effects of air pollution.

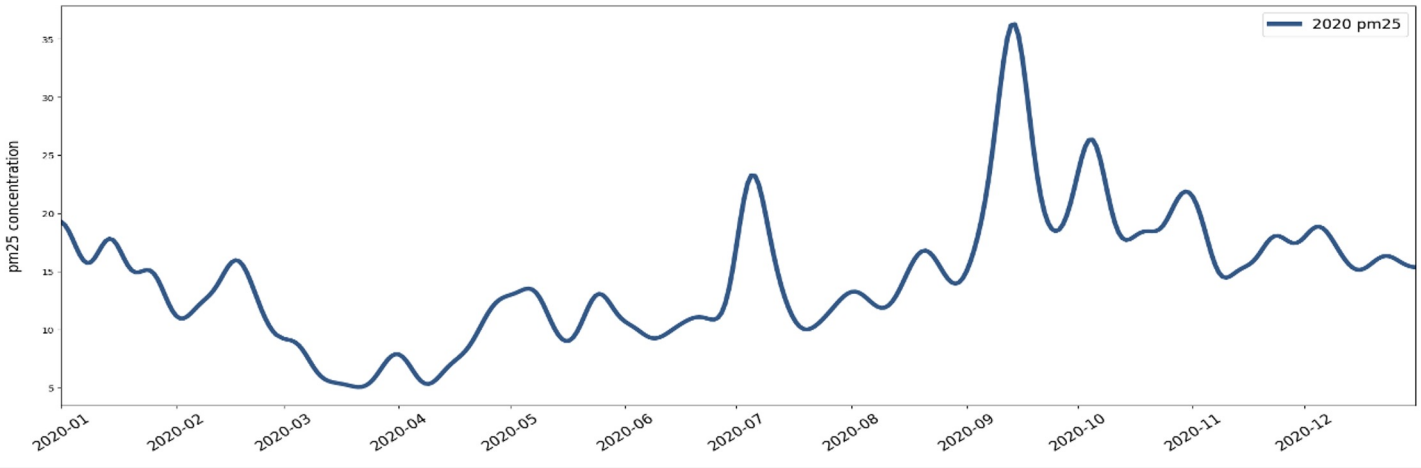
[1] UN Report 2019, UNICEF, https://news.un.org/en/story/2019/06/1039661

Air Pollution Prediction

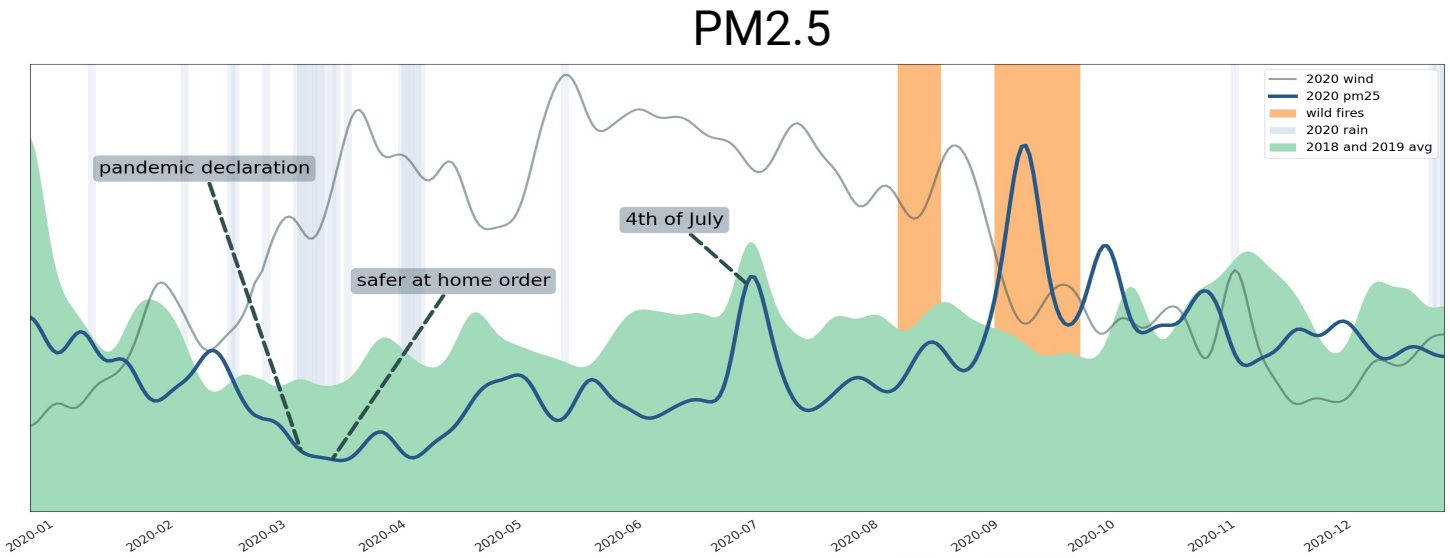
Air pollution prediction is a complex problem!

- Many factors are involved

PM2.5



Air Pollution Prediction

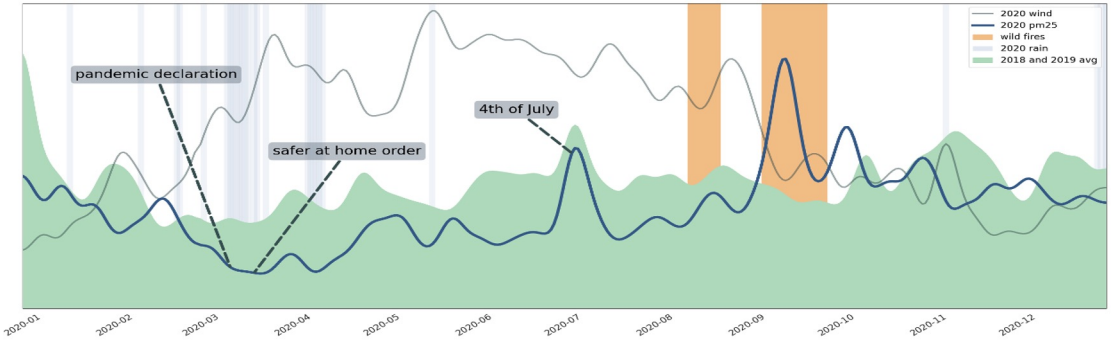


Air Pollution Prediction

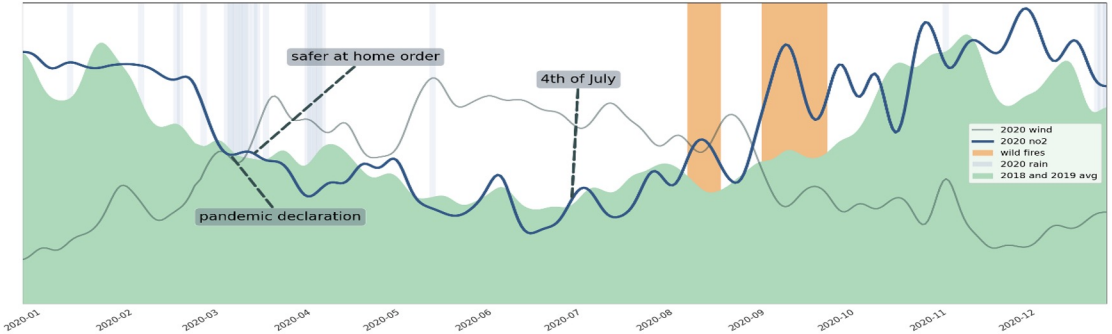
We need to:

- Take into account all factors that have impact on AQ or provide information
- Collect, process, and use data from many sources
- Have a complex machine learning model to discover, extract, and learn patterns

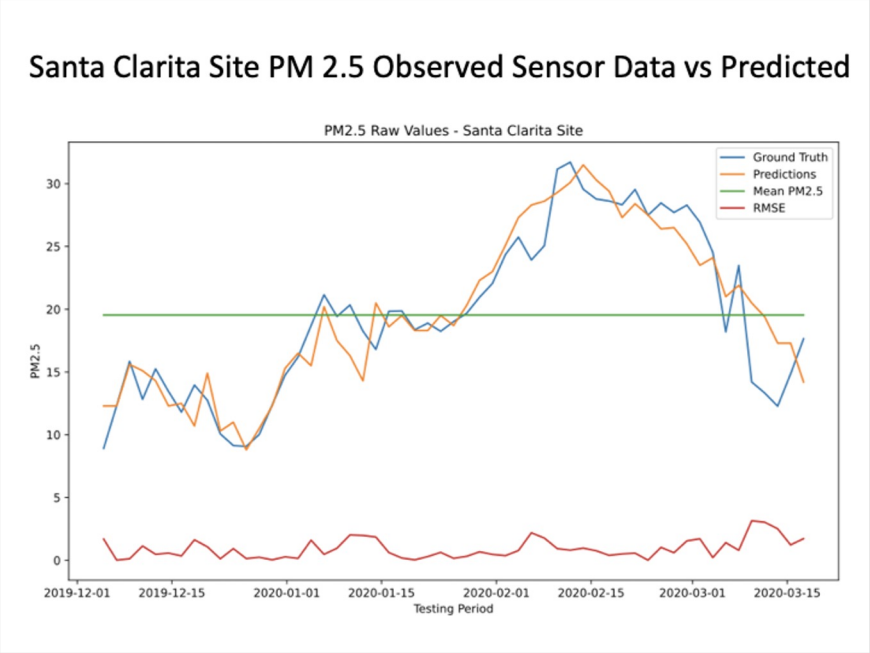
PM2.5 2020

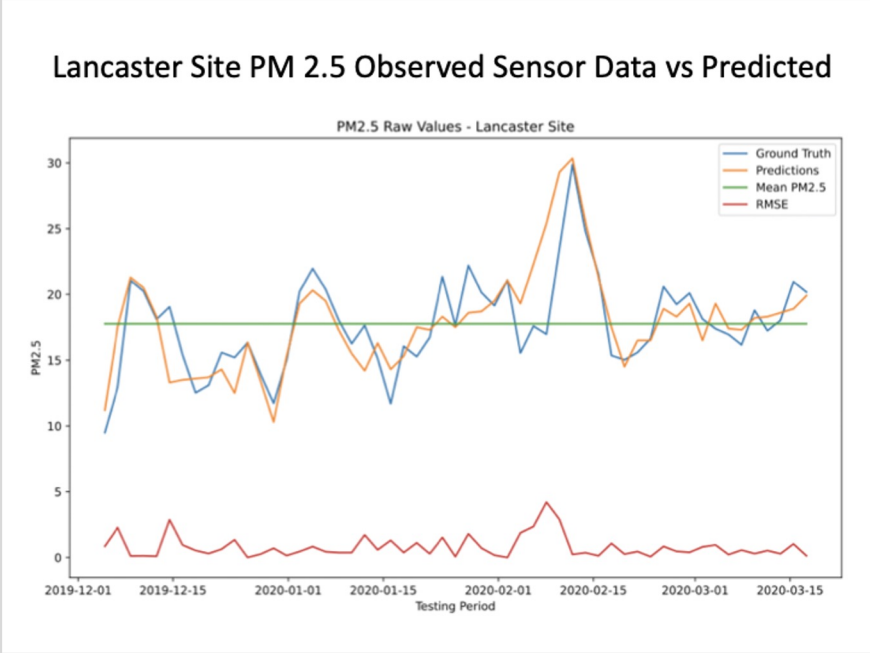


NO_2 2020



Sample Prediction Results for PM2.5 Based on Satellite Images, Ground-based Sensors and Meteorological Data



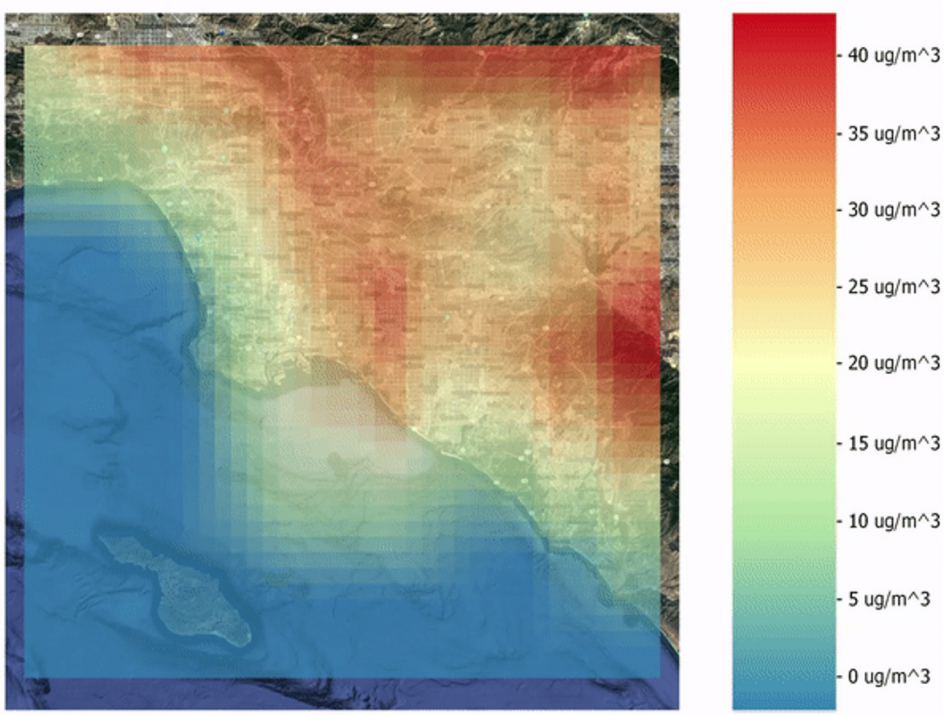


AI-based Air Quality Prediction

The average accuracy for 24-hour prediction over all site locations in LA County is 94.56%.

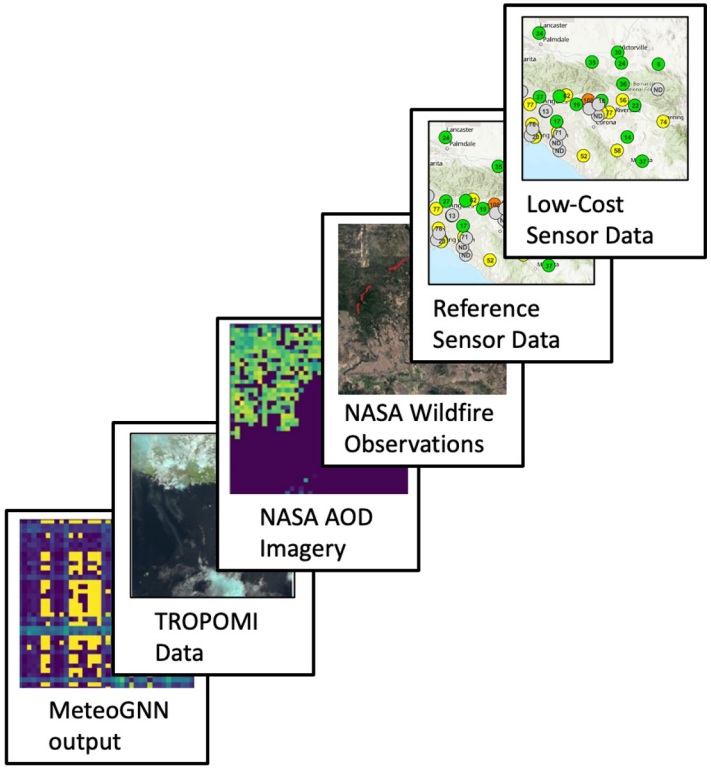
12 predictive models

PM2.5, NO2, O3, CO, CO2, SO2



- **Temporal Resolution: hourly prediction**
- **Spatial Resolution: 250 m^2**

Predictive Model and Data



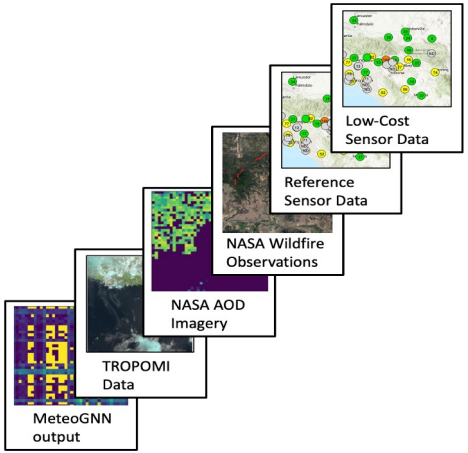
Data

- NASA AOD Imagery
- ESA/NASA TROPOMI Data
- NASA Wildfire data: MISR, MODIS, FRP
- Reference Sensors
- Low Cost Sensors: 32 AQMD sensors and 48 PurpleAir community maintained sensors.
- Meteorological data

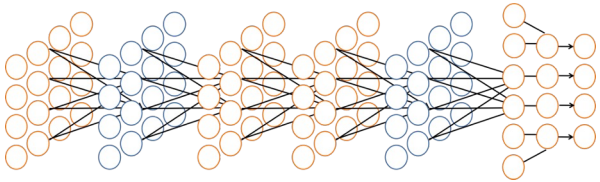
Data Processing and Data Fusion

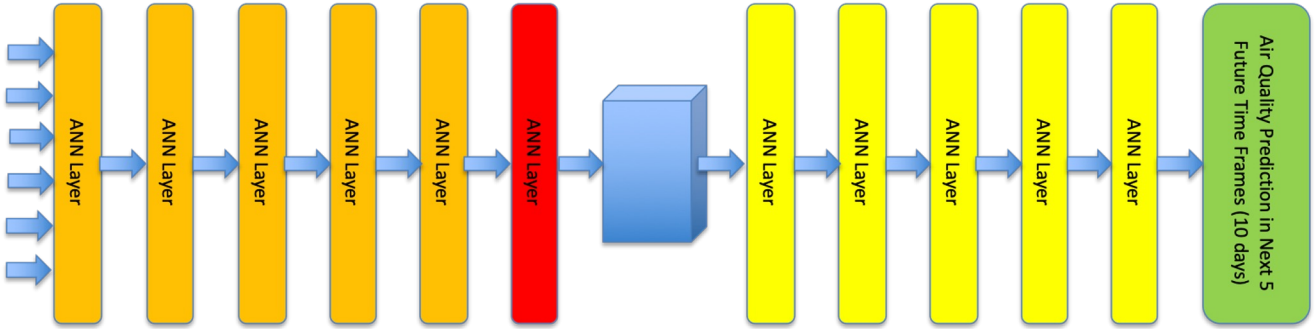
- Preprocessing and cleansing
- Outliers/trustworthiness and missing values
- Feature extraction and knowledge discovery
- Feature selection and dimensionality Reduction
- Format matching and alignments

Deep Neural Networks for Predictive Models

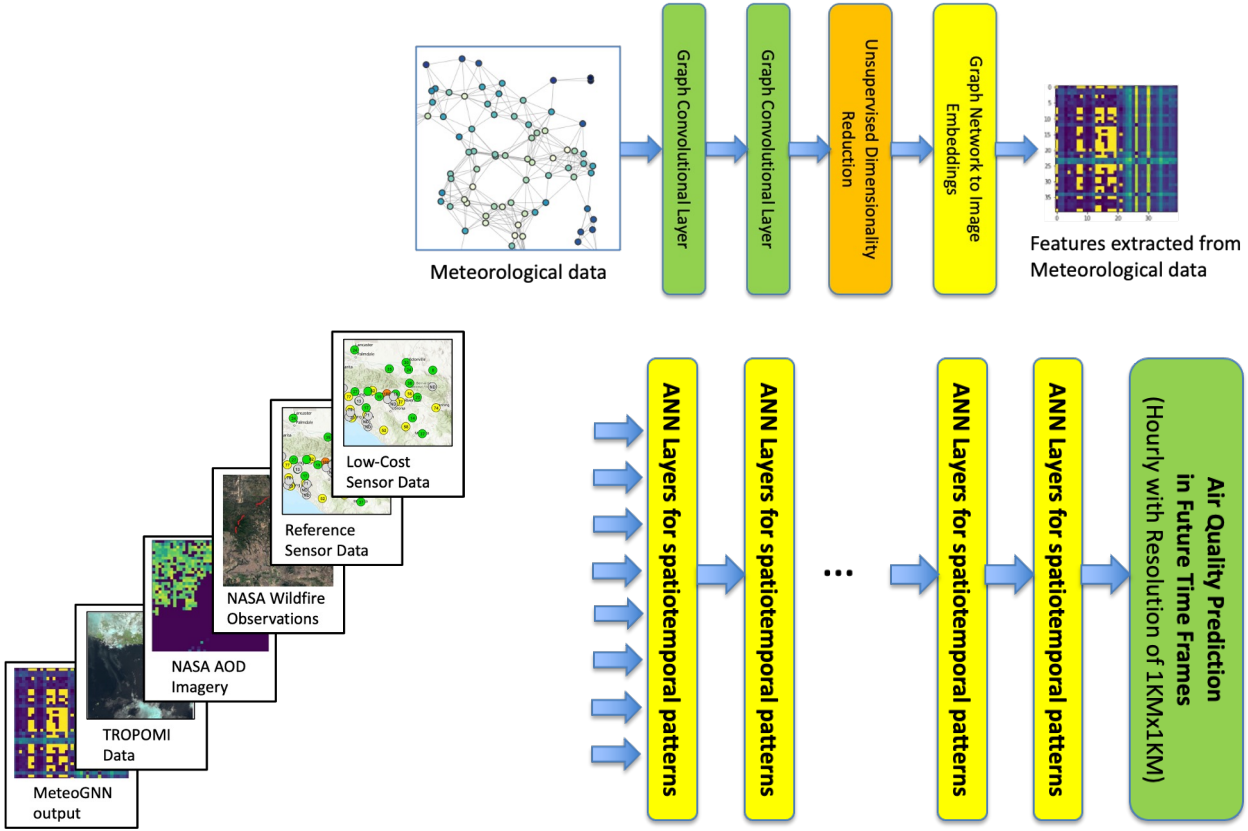


- Developed 12 Predictive models. The predictive models include several very deep neural nets to capture and learn both temporal and spatial patterns



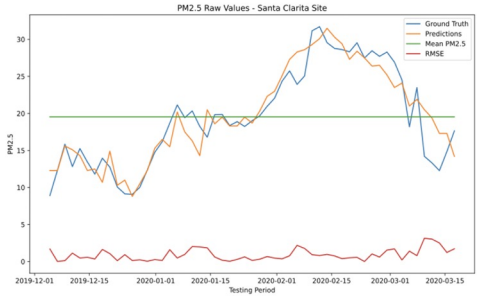


Deep Neural Networks for Predictive Models

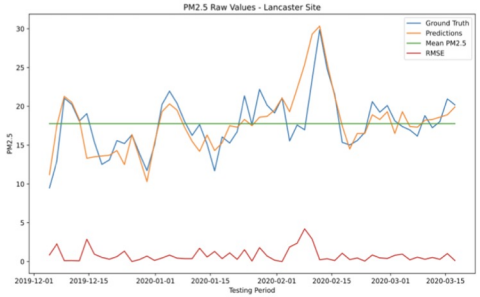


Predicting **PM2.5** Based on Satellite Observations, Ground Sensors, Meteorological Data, and Wildfire/Smoke Data

Santa Clarita Site PM 2.5 Observed Sensor Data vs Predicted



Lancaster Site PM 2.5 Observed Sensor Data vs Predicted



48-hour prediction Accuracy	Sensor Location
94%	Downtown LA
95%	Long Beach
91%	Lancaster
91%	Glendora
93%	Santa Clarita
93%	Reseda
95%	Long Beach – Rt 710

Predicting **PM2.5** Based on Satellite Observations, Ground Sensors, Meteorological Data, and Wildfire/Smoke Data

Input data

- Satellite observations NASA MODIS
- Ground-based sensors (13 in L.A. County), hourly
- Wildfire/Smoke data from NASA MODIS, MERRA-2
- Meteorological data

The average accuracy for 24-hour prediction over all site locations in LA County is 94.56%.

10-Day Prediction Accuracy	Days
93%	2 days in future
90%	4 days in future
88%	6 days in future
83%	8 days in future
80%	10 days in future

Predicting **Ozone** Based on Satellite Observations, Ground Sensors, Meteorological Data, and Wildfire/Smoke Data

48-hour prediction Accuracy	Sensor Location
93.53%	Downtown LA
95.90%	Long Beach
91.25%	Santa Clarita
88.19%	Reseda
86.23%	Lancaster
87.35%	Glendora
91.45%	Westchester
87.49%	Pico Rivera
90.04%	Compton
92.87%	Pasadena
93.10%	West LA
92.13%	Azusa
90.59%	Pomona

Accuracy	Frame #
91%	Frame 1: 2 days in future
89%	Frame 2: 4 days in future
86%	Frame 3: 6 days in future
84%	Frame 4: 8 days in future
80%	Frame 5: 10 days in future

- Satellite observations NASA MODIS
- Ground-based sensors (13 in L.A. County), hourly
- Wildfire/Smoke data from NASA MODIS, MERRA-2
- Meteorological data

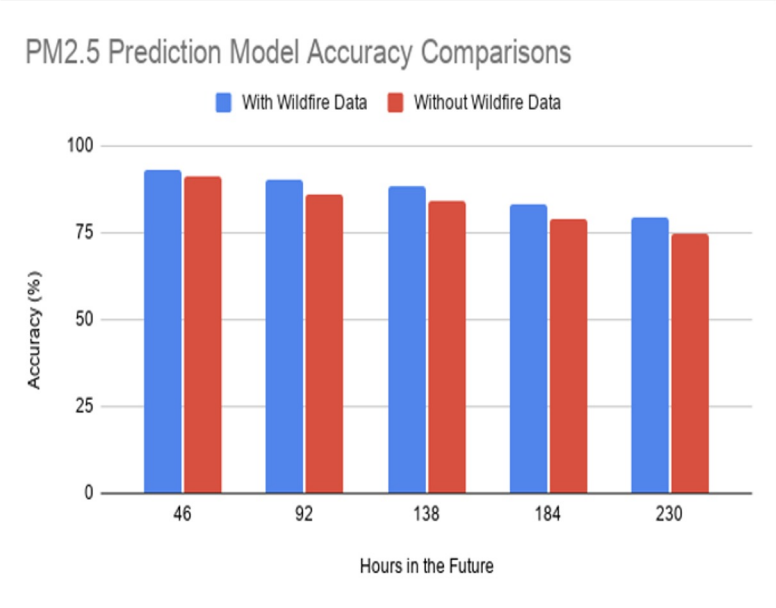
Predicting **NO_2** Based on Satellite Observations, Ground Sensors, Meteorological Data, and Wildfire/Smoke Data

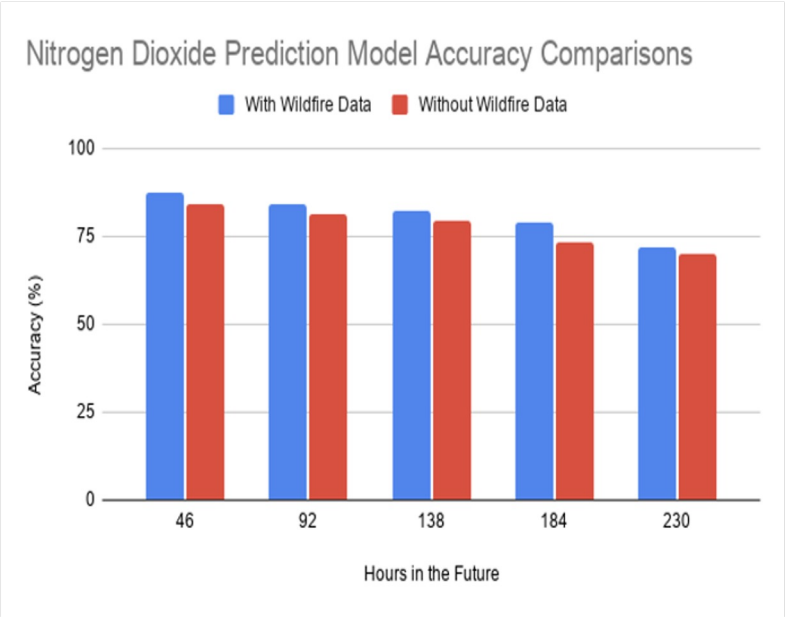
24-hour prediction Accuracy	Sensor Location
93%	Downtown LA
91%	Long Beach
91%	Santa Clarita
89%	Reseda
87%	Lancaster
88%	Glendora
91%	Westchester
91%	Pico Rivera
95%	Compton
92%	Pasadena
90%	West LA
92%	Azusa
92%	Pomona

Accuracy	Frame #
87.62%	Frame 1: 2 days in future
84.15%	Frame 2: 4 days in future
82.38%	Frame 3: 6 days in future
79.06%	Frame 4: 8 days in future
72%	Frame 5: 10 days in future

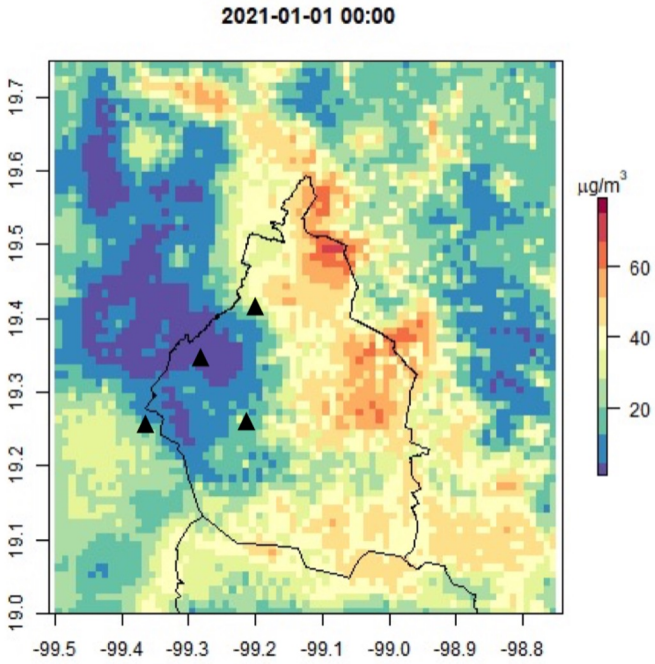
- Satellite observations NASA MODIS
- Ground-based sensors (13 in L.A. County), hourly
- Wildfire/Smoke data from NASA MODIS, MERRA-2
- Meteorological data

Model Comparisons: Effect of Wildfire/Smoke Data





PWWB forecast in Mexico City for $PM_{2.5}$



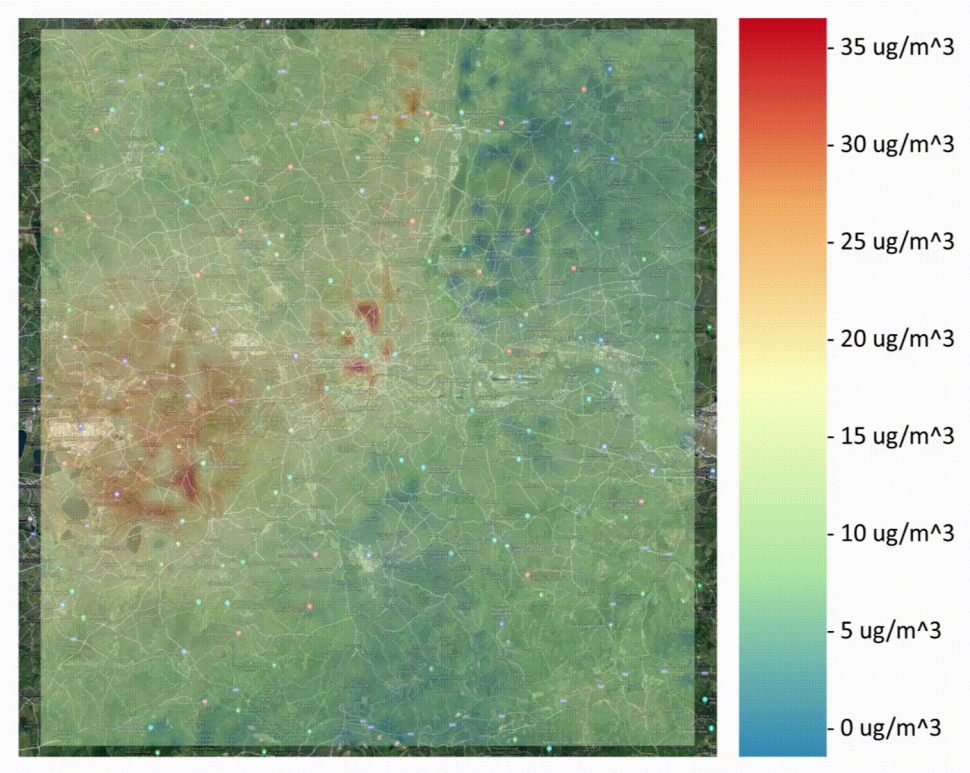
▲ Top 4 monitoring stations

Prediction accuracy Nov 1 to Dec 31

Average Accuracy	Monitoring Station
94.87%	**Santa Fe**
94.72%	**Ajusco Medio***
94.53%	**Miguel Hidalgo***
94.51%	**Investigaciones Nucleares**
94.42%	Hospital General de México*
94.39%	Benito Juárez
94.34%	Tlalnepantla
94.19%	San Agustín
94.18%	Merced
94.17%	Gustavo A. Madero
93.95%	Ajusco
93.69%	Nezahualcóyotl
93.68%	Centro de Ciencias de la Atmósfera
93.56%	Xalostoc*
93.39%	UAM Xochimilco

London, U.K.

Hourly prediction



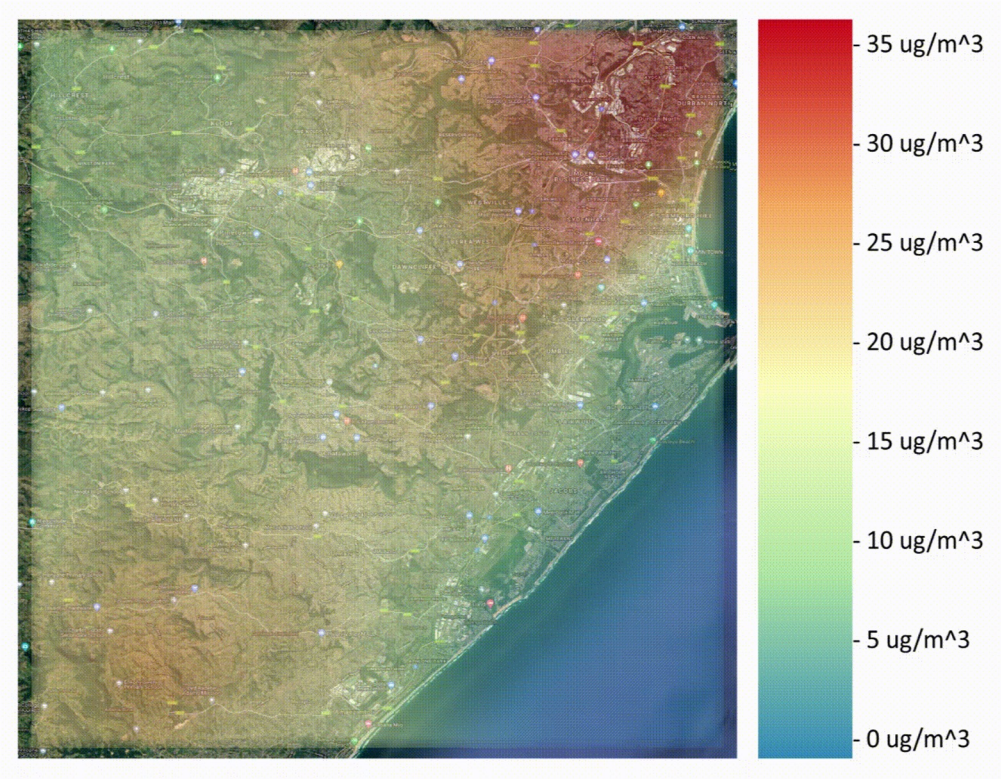
London Predicted PM2.5 Results

Predicted ground-level PM2.5 in 12 sensor locations in London hourly

Average Accuracy	Sensor Location
91.03%	London Teddington Bushy Park
91.14%	Kensington and Chelsea
91.04%	Sutton - Beddington Lane
89.61%	Camden - Bloomsbury
91.17%	City of London - Farringdon Street
90.90%	City of London - The Aldgate School
90.77%	Tower Hamlets - Blackwall
90.65%	Greenwich - Westhorne Avenue
91.03%	Greenwich - A206 Burrage Grove
90.92%	Greenwich - Plumstead High Street
90.98%	Greenwich - Falconwood FDMS
91.02%	Havering - Rainham

Durban, South Africa

Hourly prediction



Durban Predicted PM2.5 Results

Predicted ground-level PM2.5 in 4 sensor locations in Durban hourly

Average Accuracy	Sensor Location
94.54%	Settlers
94.53%	Wentworth
93.88%	Durban City Hall
95.61%	New Germany

Thank you!

More info:
Mohammad Pourhomayoun
mpourho@calstatela.edu

airquality.lacity.org
www.ai-agora.com